JC18 Rec'd PCT/PTO 20 JUN 2001

BAKER BOTTS LLP TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35.U.S.C. 371	EXPRESS MAIL LABEL No. EF321684975US	DATE June 20, 2001
	ATTORNEY'S DOCKET NO A34274-PCT-USA	
	U S APPLICATION NO 09/869086 (To be assigned)	

INTERNATIONAL APPLICATION NO	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/GB99/04007	December 1, 1999	December 22, 1998

TITLE OF INVENTION DISTRIBUTED HIERARCHICAL SCHEDULING AND ARBITRATION FOR BANDWIDTH ALLOCATION

APPLICANT(S) FOR DO/EO/US PIEKARSKI, Marek Stephen; JOHNSON, Ian David

Applicant herewith submits to the United States Designated /Elected Office (DO/EO/US) the following items and other information:

1. [✓] This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. [ ] This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(I).
4. [✓] A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. [✓] A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. [ ] is transmitted herewith (required only if not transmitted by the International Bureau).
   b. [✓] has been transmitted by the International Bureau.
   c. [ ] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. [✓] A copy of the International Search Report (PCT/ISA/210)
   a. [ ] are transmitted herewith (required only if not transmitted by the International Bureau).
   b. [✓] have been transmitted by the International Bureau
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.
8. [ ] A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. [✓] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. [ ] A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern other document(s) or information included:

11. [ ] A copy of the International Preliminary Examination Report (PCT/IPEA/409)
12. [✓] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [✓] A FIRST preliminary amendment.
    [ ] A SECOND or SUBSEQUENT preliminary amendment.
14. [ ] A substitute specification.
15. [ ] A change of power of attorney and/or address letter.
16. [ ] Other items or information:
    a. [ ] a copy of the International Search Report (PCT/ISA/210)
    b. [ ] a copy of the International Preliminary Examination Report (PCT/IPEA/409)

Drawings (Figs. 1-7), Verified Statement Claiming Small Entity Status, Check for $740.00, Postcard

09/869086 Rec'd PCT/PTO 20 JUN 2001

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/GB99/04007	December 1, 1999	December 22, 1998

17. [ ] The following fees are submitted: — CALCULATIONS — PTO USE ONLY

Basic National Fee (37 CFR 1.492(a)(1)-(5):

Neither international preliminary examination fee (37 CFR 1.482)

Nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO (1.492(a)(3)) $1,000.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO (1.492(a)(5) $860.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO(1.492(a)(2)) $710.00

International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) (1.492(a)(1)) $690.00

International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) $ 100.00

				Calculations	PTO use only
ENTER APPROPRIATE BASIC FEE AMOUNT =				$ 1,000	
Surcharge of $130.00 for furnishing the oath or declaration later than [✓] 20 [ ] 30 months from the earliest claimed priority date (37 C.F.R. 1.492)(e)).				$ 130	
Claims	**Number Filed**	**Number Extra**	**Rate**	$	
Total Claims	14 -20=	0	X $ 18.00	$ 0	
Independent Claims	2 -3=	0	X $ 80.00	$ 0	
Multiple dependent claim(s) (if applicable)			+ $270.00	$ 270	
TOTAL OF ABOVE CALCULATIONS =				$ 1,400	
Reduction by ½ for filing by small entity, if applicable.				$ 700	
SUBTOTAL =				$ 700	
Processing fee of $130.00 for furnishing the English translation later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(f)). +				$	
TOTAL NATIONAL FEE =				$ 700	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property +				$ 40	
TOTAL FEES ENCLOSED =				$ 740	
				Amt. refunded	$
				charged	$

a. [✓] A check in the amount of $ 740.00 to cover the above fees is enclosed.

b. [ ] Please charge our Deposit Account No. 02-4377 in amount of $________ to cover the above fees. A copy of this sheet is enclosed.

c. [✓] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 02-4377. A copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Paul A. Ragusa
BAKER BOTTS L.L.P.
30 Rockefeller Plaza
New York, New York 10112-4498

Attorney: Paul A. Ragusa PTO Reg: 38,587

June 20, 2001

Date

09/869086

20 JUN 2001

BAKER BOTTS LLP

Attorney Docket Number: A34274-PCT-USA

Title: DISTRIBUTED HIERARCHICAL SCHEDULING AND ARBITRATION FOR BANDWIDTH ALLOCATION

Use Space Below for Additional Information:

A34274-PCT-USA - 072854.0119

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	:	Piekarski et al.
Serial No.	:	TBD
Filed	:	June 20, 2001
For	:	DISTRIBUTED HIERARCHICAL SCHEDULING AND ARBITRATION FOR BANDWIDTH ALLOCATION

PRELIMINARY AMENDMENT

I hereby certify that this paper is being deposited with the United States Postal Service as Express Mail in an envelope addressed to: Assistant Commissioner for Patents, Washington, D.C. 20231

June 20, 2001
Date of Deposit

Paul A. Ragusa
Attorney Name

38,587
PTO Registration No.

Signature

June 20, 2001
Date of Signature

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Preliminary to examination, please amend the claims as follows.

IN THE CLAIMS:

Please amend claims 1-14.

1. (Amended) A method of scheduling the passage of data cells from M low-bandwidth data sources to M low-bandwidth data destinations in a data switching apparatus which includes:

M/N ingress multiplexers, each arranged to receive data cells from a respective set of N said low-bandwidth data sources,

M/N egress multiplexers, each arranged to transmit data cells to a respective set of N said low-bandwidth data destinations],

a master control unit, and

a central switch having M/N high-bandwidth input ports arranged to receive data cells from respective said ingress multiplexers, and M/N high-bandwidth output ports arranged to transmit data cells to respective said egress multiplexers, the central switch selectively interconnecting the input ports and output ports, under the direction of the master control unit,

the method comprising the steps of:

maintaining N input queues in each of said ingress multiplexers for queuing data cells received from the N respective said data sources, and maintaining M virtual output queues for queuing data cells directed to respective said data destinations;

maintaining a respective ingress port table in each of said ingress multiplexers, each ingress port table having NxM entries, each entry corresponding to a respective combination of a said data source for that ingress port and a said data destination,

transferring data cells from said input queues to said virtual output queues by each of said ingress multiplexers with a relative frequency according to value of the corresponding entry of the ingress port table;

maintaining a respective egress port table in each of said ingress multiplexers, the egress port table having M entries, each entry corresponding to a respective said data destination,

transferring data cells from said virtual output queues to said respective input ports of the central switch by each of said ingress multiplexers with a relative frequency according to the value of the corresponding entry of the egress port table;

maintaining a central allocation table in the master control unit having $(M/N)^2$ entries, each corresponding to a respective combination of an input port and an output port, and

controlling the central switch by the master control unit to interconnect pairs of said input ports and output ports with a relative frequency according to the value of the corresponding entry of the central allocation table

whereby said ingress port tables, egress port tables and central allocation table together determine the bandwidth through the digital data switching apparatus from each said data source to each said data destination

2. (Amended) The method according to claim 1 wherein each said ingress multiplexer, for each virtual output queue, transfers data cells to that virtual output queue from said input queues in accordance with a N-way weighted round robin, using N weights determined respectively by the N entries of the ingress port table for that virtual output queue.

3. (Amended) The method according to claim 2, wherein each weight is defined by a number of bits w, and the N-way weighted round robin for each virtual output queue is

implemented by an $N(2^w-1)$-way unweighted round robin using a request vector list constructed by interleaving N words of (2^w-1) bits each, each word corresponding to a respective input queue and having a number of "1"s determined by the entry of the ingress port table for that input queue and that virtual output queue.

4. (Amended) The method according to claim 3, wherein the request vector list is separated into a plurality of round robin blocks, each corresponding to a respective input queue, a first round robin process being performed independently within each block, and a second round robin process being performed to make a selection among the blocks.

5. (Amended) The method according to claim 1, wherein the ingress port table the egress port table and the central allocation table re all programmed from an external source.

6. (Amended) The method according to claim 5, wherein the external source uses parameters characterizing the length of each virtual output queue and the urgency of each virtual output queue.

7. (Amended) The method according to claim 6, wherein the external source uses a set of sensitivities relating to length, urgency and pseudo-static bandwidth allocation.

8. (Amended) A digital data switching apparatus for transmitting data from M low-bandwidth data sources to M low-bandwidth data destinations, comprising:

M/N ingress multiplexers for receiving data cells from respective sets of N said low-bandwidth data sources,

M/N egress multiplexers for transmitting data cells to respective sets of N said low-bandwidth data destinations,

a master control unit,

a central switch having M/N high-bandwidth input ports arranged to receive data cells from respective said ingress multiplexers, and M/N high-bandwidth output ports arranged to transmit data cells to respective said egress multiplexers, the central switch being arranged selectively to interconnect the input ports and output ports, under the direction of the master control unit,

each said ingress multiplexer being arranged to maintain N input queues for queuing data cells received from respective said data sources, and to maintain M virtual output queues for queuing data cells directed to respective said data destinations;

wherein each ingress multiplexer is arranged to maintain a respective ingress port table, each ingress port table having NxM entries, each entry corresponding to a respective combination of a said data source and a said data destination, and each ingress multiplexer is arranged to transfer data cells from said input queues to said virtual output queues with a relative frequency according to value of the corresponding entry of the ingress port table;

each ingress multiplexer is arranged to maintain a respective egress port table, the egress port table having M entries, each corresponding to a respective said data destination, and each ingress multiplier is arranged to transfer data cells from said virtual output queues to said respective input ports of the central switch with a relative frequency according to value of the corresponding entry of the egress port table,

and the master control unit is arranged to maintain a central allocation table having $(M/N)^2$ entries, each corresponding to a respective combination of an input port and an output port, and the master control unit controls the central switch to interconnect pairs of said input ports and output ports with a relative frequency according to the value of the corresponding

entry of the central allocation table;

whereby said ingress port tables, egress port tables and central allocation table together determine the bandwidth through the digital data switching apparatus from each said data source to each said data destination.

9. (Amended) The apparatus according to claim 8, wherein each said ingress multiplexer is arranged, for each virtual output queue, to transfer data cells to that virtual output queue from said input queues in accordance with a N-way weighted round robin, using N weights determined respectively by the N entries of the ingress port table for that virtual output queue.

10. (Amended) The apparatus according to claim 8, wherein each weight has a number of bits w, and the N-way weighted round robin for each virtual output queue is implemented by an $N(2^w-1)$- way unweighted round robin using a request vector list constructed by interleaving N words of (2^w-1) bits each, each word corresponding to a respective input queue and having a number of "1"s determined by the entry of the ingress port table for that input queue and that virtual output queue.

11. (Amended) The apparatus according to claim 10, wherein the request vector list is separated into a plurality of round robin blocks, each corresponding to a respective input queue, each ingress multiplexer being arranged to preform a first round robin process independently within each block, and a second round robin process to make a selection among the blocks.

12. (Amended) The apparatus according to claim 8, further comprising an external source unit arranged to program the ingress port table, the egress port table and the central allocation table.

13. (Amended) The apparatus according to claim 12, wherein the external source unit is arranged to operate using parameters characterizing the length of the virtual output queue and the urgency of the virtual output queue.

14. (Amended) The apparatus according to claim 13, wherein the external source unit is arranged to operate using a set of sensitivities relating to the length, urgency and pseudo-static bandwidth allocation.

REMARKS

By the foregoing amendment, claims 1-14 have been amended to remove multiple dependencies and to place the claims in proper U.S. format. No new matter has been added.

Respectfully submitted,

Paul A. Ragusa
Patent Office Reg. No. 38,587

Attorney for Applicants

BAKER BOTTS L.L.P.
30 Rockefeller Plaza
New York, NY 10112
(212) 408-2588

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS:

Please amend claims 1-14.

1. (Amended) A method of scheduling the passage of data cells from M low-bandwidth data sources [(4)] to M low-bandwidth data destinations [(5), said method being performed by] in a data switching apparatus [including:] which includes:

M/N ingress multiplexers [(2; 71, 72)], each arranged to receive data cells from a respective set of N said low-bandwidth data sources [(4)],

M/N egress multiplexers [(3; 74, 75)], each arranged to transmit data cells to a respective set of N said low-bandwidth data destinations [(5)],

a master control unit, and

a central switch having M/N high-bandwidth input ports arranged to receive data cells from respective said ingress multiplexers [(2; 71, 72)], and M/N high-bandwidth output ports arranged to transmit data cells to respective said egress multiplexers [(3; 74, 75)], the central switch selectively interconnecting the input ports and output ports, under the direction of the master control unit,

the method [including:] comprising the steps of:

[each said ingress multiplexer (2; 71, 72)] maintaining N input queues [(22)] in each of said ingress multiplexers for queuing data cells received from the N respective said data sources, and maintaining M virtual output queues [(24)] for queuing data cells directed to respective said data destinations;

[and the method being characterized in that:]

[each ingress miltiplexer further] maintain[s]ing a respective ingress port table [(27; 77)] in each of said ingress multiplexers, each ingress port table having NxM entries, each entry corresponding to a respective combination of a said data source for that ingress port and a said data destination,

[each ingress multiplexer] transfer[s]ring data cells from said input queues to said virtual output queues by each of said ingress multiplexers with a relative frequency according to value of the corresponding entry of the ingress port table [(27; 77)];

[each ingress multiplexer further] maintain[s]ing a respective egress port table [(28; 78)] in each of said ingress multiplexers, the egress port table having M entries, each entry corresponding to a respective said data destination,

[each ingress multiplexer] transfer[s]ring data cells from said virtual output queues to said respective input ports of the central switch by each of said ingress multiplexers with a relative frequency according to the value of the corresponding entry of the egress port table [(28; 78)];

[the master control unit] maintain[s]ing a central allocation table [(79)] in the master control unit having $(M/N)^2$ entries, each corresponding to a respective combination of an input port and an output port, and

[the master control unit] control[s]ling the central switch by the master control unit to interconnect pairs of said input ports and output ports with a relative frequency according to the value of the corresponding entry of the central allocation table [(79)];

whereby said ingress port tables [(27; 77)], egress port tables [(28; 78)] and

central allocation table [(79)] together determine the bandwidth through the digital data switching apparatus from each said data source [(4)] to each said data destination [(5)].

2. (Amended) The [A] method according to claim 1 wherein [in which] each said ingress multiplexer [(2; 71, 72)], for each virtual output queue (24), transfers data cells to that virtual output queue from said input queues in accordance with a N-way weighted round robin, using N weights determined respectively by the N entries of the ingress port table for that virtual output queue.

3. (Amended) The [A] method according to claim 2, wherein [in which] each weight is defined by a number of bits w, and the N-way weighted round robin for each virtual output queue is implemented by an $N(2^w-1)$-way unweighted round robin using a request vector list constructed by interleaving N words of (2^w-1) bits each, each word corresponding to a respective input queue and having a number of "1"s determined by the entry of the ingress port table for that input queue and that virtual output queue.

4. (Amended) The [A] method according to claim 3, wherein [in which] the request vector list is separated into a plurality of round robin blocks, each corresponding to a respective input queue, a first round robin process being performed independently within each block, and a second round robin process being performed to make a selection among the blocks.

5. (Amended) The [A] method according to [any preceding] claim 1, wherein [in which] the ingress port table [(27; 77)], the egress port table [(28; 78)] and the central allocation table [(79)] are all programmed from an external source.

6. (Amended) The [A] method according to claim 5, wherein [in which] the external source uses parameters characterizing the length of each virtual output queue and the urgency of each virtual output queue.

7. (Amended) The [A] method according to claim 6, wherein [in which] the external source uses a set of sensitivities relating to length, urgency and pseudo-static bandwidth allocation.

8. (Amended) A digital data switching apparatus for transmitting data from M low-bandwidth data sources [(4)] to M low-bandwidth data destinations [(5)], [the apparatus including] comprising:

M/N ingress multiplexers [(2; 71, 72)] for receiving data cells from respective sets of N said low-bandwidth data sources [(4)],

M/N egress multiplexers [(3; 74, 75)] for transmitting data cells to respective sets of N said low-bandwidth data destinations [(5)],

a master control unit, [and]

a central switch having M/N high-bandwidth input ports arranged to receive data cells from respective said ingress multiplexers [(2; 71, 72)], and M/N high-bandwidth output ports arranged to transmit data cells to respective said egress multiplexers [(3; 74, 75)], the central switch being arranged selectively to interconnect the input ports and output ports, under the direction of the master control unit,

each said ingress multiplexer [(2; 71, 72)] being arranged to maintain N input queues [(22)] for queuing data cells received from respective said data sources, and to maintain

M virtual output queues [(24)] for queuing data cells directed to respective said data destinations;

[characterized in that]:

wherein each ingress multiplexer [(2; 71, 72)] is arranged to maintain a respective ingress port table [(27; 77)], each ingress port table having NxM entries, each entry corresponding to a respective combination of a said data source and a said data destination, and each ingress multiplexer is arranged to transfer data cells from said input queues to said virtual output queues with a relative frequency according to value of the corresponding entry of the ingress port table;

each ingress multiplexer is arranged to maintain a respective egress port table [(28; 78)], the egress port table having M entries, each corresponding to a respective said data destination [(5)], and each ingress multiplier is arranged to transfer data cells from said virtual output queues to said respective input ports of the central switch with a relative frequency according to value of the corresponding entry of the egress port table [(28; 78)],

and the master control unit is arranged to maintain a central allocation table [(79)] having $(M/N)^2$ entries, each corresponding to a respective combination of an input port and an output port, and the master control unit controls the central switch to interconnect pairs of said input ports and output ports with a relative frequency according to the value of the corresponding entry of the central allocation table [(79)];

whereby said ingress port tables [(27; 77)], egress port tables [(28; 78)] and central allocation table [(79)] together determine the bandwidth through the digital data switching apparatus from each said data source [(4)] to each said data destination [(5)].

9. (Amended) The [An] apparatus according to claim [1] 8, wherein [in which] each said ingress multiplexer is arranged, for each virtual output queue [(24)], to transfer data cells to that virtual output queue from said input queues [(22)] in accordance with a N-way weighted round robin, using N weights determined respectively by the N entries of the ingress port table [(27; 77)] for that virtual output queue.

10. (Amended) The [An] apparatus according to claim [11] 8, wherein [in which] each weight has a number of bits w, and the N-way weighted round robin for each virtual output queue is implemented by an $N(2^w-1)$- way unweighted round robin using a request vector list constructed by interleaving N words of (2^w-1) bits each, each word corresponding to a respective input queue and having a number of "1"s determined by the entry of the ingress port table for that input queue and that virtual output queue.

11. (Amended) The [An] apparatus according to claim 10, wherein [in which] the request vector list is separated into a plurality of round robin blocks, each corresponding to a respective input queue, each ingress multiplexer being arranged to preform a first round robin process independently within each block, and a second round robin process to make a selection among the blocks.

12. (Amended) The [An] apparatus according to [any preceding] claim 8, further comprising an external source unit arranged to program the ingress port table [(27; 77)], the egress port table [(28; 78)] and the central allocation table [(29)].

13. (Amended) The [An] apparatus according to claim 12, wherein [in which] the external source unit is arranged to operate using parameters characterizing the length of the virtual output queue and the urgency of the virtual output queue.

14. (Amended) The [An] apparatus according to claim 13, wherein [in which] the external source unit is arranged to operate using a set of sensitivities relating to the length, urgency and pseudo-static bandwidth allocation.

09/869086

7/PRTS

FILE NO.: A34274 - 072854.0119

JC18 Rec'd PCT/PTC 20 JUN 2001

BAKER BOTTS L.L.P
30 ROCKEFELLER PLAZA
NEW YORK, NEW YORK 10112

TO ALL WHOM IT MAY CONCERN:

Be it known that We, Marek Stephen Piekarski and Ian David Johnson, respectively, citizens of Great Britain, whose post office addresses are 20 Gawsworth road, Macclesfield, Cheshire SK11 8UE, Great Britain and 11 Seel Street, Mosley OL5 0EW, Great Britain, respectively, have made an invention in

DISTRIBUTED HIERARCHICAL SCHEDULING AND ARBITRATION FOR BANDWIDTH ALLOCATION

of which the following is a

SPECIFICATION

BACKGROUND OF THE INVENTION

[0001] The present invention relates to data switching systems and is more particularly concerned with the scheduling and arbitration arrangements for such systems.

[0002] The continual growth of demand for manageable bandwidth in networks requires the development of new techniques in switch design which decouples the complexity of control from the scale of the port count and aggregate bandwidth. This invention describes a switch architecture and a set of methods which provide the means by which switches of arbitrary size may be constructed whilst maintaining the ability to allocate guaranteed bandwidth to each possible connection through the switch. A digital switch is used to route data streams from a set of source components to a set of destination components. A cell-based switch operates on data which is packetized into streams of equal size cells. In a large switch the routing functions may

be implemented hierarchically, that is sets of lower bandwidth ports are aggregated into a smaller number of higher bandwidth ports which are then interconnected in a central switch.

[0003] It is an object of the present invention to provide a bandwidth allocation arrangement which may be used in such a hierarchical switch.

SUMMARY OF THE INVENTION

[0004] According to the present invention there is provided a scheduling and arbitration process for use in a digital data switching arrangement of the type in which a central switch under the direction of a master control provides the cross-connections between a number of high-bandwidth ports to which are connected on the ingress side of the central switch a number of ingress multiplexers, one for each high-bandwidth input port and on the egress side a number of egress multiplexers, one for each high-bandwidth output port, each ingress multiplexer including a set of N input queues serving N low-bandwidth data sources and a set of M virtual output queues serving M low-bandwidth output data sources, characterized in that the scheduling and arbitration arrangement includes three bandwidth allocation tables, an ingress port table associated with the input queues and having NxM entries each arranged to define the bandwidth for a particular virtual output queue, an egress port table associated with the virtual output queues and having M entries each arranged to define the bandwidth allocation of a high-bandwidth port of the central switch to a virtual output queue and a central allocation table located in the master control and having $(MxN)^2$ entries each of which specifies the weights allocated to each possible connection through the central switch.

[0005] According to a feature of the invention there is provided a scheduling and arbitration process in which the scheduling of the input queues is performed in accordance with an N-way weighted round robin.

[0006] According to a further feature of the invention there is provided an implementation of the N-way weighted round robin by an $N.(2^{w}-1)$-way unweighted round robin where u, is the number of bits defining a weight using a list constructed by interleaving N words of $(2^{w}-1)$ bits each, with w_n 1's in a word, where w_n is the weight of the queue n.

BRIEF DESCRIPTION OF THE DRAWING

[0007] The invention, together with its various features, will be more readily understood from the following description of one embodiment, which should be read in conjunction with the accompanying drawings. In the drawings:

[0008] Figure 1 shows a simplified form of a data switch, Figure 2 shows an egress multiplexer;

[0009] Figure 3 shows the weighted round robin arbiter for use in the egress multiplexer;

[0010] Figure 4 shows the partitioning of the round robin arbiter, Figure 5 shows the operation of the round robin arbiter;

[0011] Figure 6 shows the allocations for a 4-port interconnect with 3 bit weights; and

[0012] Figure 7 shows a block diagram of a small switch based on the principles of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0013] Referring now to Figure 1, this shows a schematic diagram of a hierarchical switch. The central interconnect 1 provides the cross-connections between a number of high-bandwidth ports. A set of multiplexers 2 on the ingress side and demultiplexers 3 on the egress side provides the aggregation function between the low and high-bandwidth ports. The low bandwidth ports provide connections from the switch to the data sources 4 on the ingress side

and the data destinations 5 on the egress side. In practice, a switch is required to support full-duplex ports, so that an ingress multiplexer and its corresponding demultiplexer may be considered a single full-duplex device which will be hereafter termed a "router" Typically the data switch may be of the type disclosed in out co-pending patent application No. PCT/GB99/03748.

[0014] It should be noted that the central interconnect 1 may itself be a hierarchical switch, that is the methods described may be applied to switches with an arbitrary number of hierarchical levels.

[0015] The aim of these methods is to provide a mechanism whereby the data stream from the switch to a particular destination, which comprises a sequence of cells interleaved from various data sources, may be controlled such that predetermined proportions of its bandwidth are guaranteed to cells from each data source.

[0016] Figure 2 shows the architecture of an ingress multiplexer. An ingress multiplexer receives a set of data streams from the data sources via a set of low-bandwidth input ports. Each data stream is a sequence of equal size cells (that is, an equal number of bits of data). A set of N low-bandwidth ports 21 each fills one of the N input queues 22. An ingress control unit 23 extracts the destination address from each of the cells in the input queues and transfers them into a set of M virtual output queues 24. There is one virtual output queue for each low-bandwidth output port in the switch. The ingress multiplexer also contains an interconnect link control unit 25 which implements this function by scheduling cells from the virtual output queues 24 across the high-bandwidth link 26 to the central interconnect 1 according to an M-entry egress table 28.

[0017] In addition to the data flow indicated by the arrows in Figure 1, there is also a flow of backpressure or flow-control information associated with each of the data flows. This

control flow is indicated in Figure 2 by dashed arrows. The ingress multiplexer contains an NxM-entry ingress port table 27, which defines how its bandwidth to a particular egress port (via a particular virtual output queue) is distributed across the input ports. This table is used by the ingress control unit 23 to determine when (and to what degree) to exert backpressure to the data source resolved down to an individual virtual output queue.

[0018] The ingress multiplexer 2 of Figure 1 sends control information to the central interconnect I indicating the state of the virtual output queues in the form of "connection requests". The central interconnect responds with a sequence of connections which it will establish between the ingress and egress routers. These are "connection grants". The ingress multiplexer 2 must now allocate the bandwidth to each egress demultiplexer 3 provided by the central interconnect 1 across the virtual output queues associated with each egress demultiplexer.

[0019] The deterministic scheduling function of the interconnect link control unit 25 may be defined as a weighted round robin (WRR) arbiter. The interconnect link control unit 25 receives a connection grant to a particular egress demultiplexer 3 from the central interconnect l and must select one of the N virtual output queues associated with that egress demultiplexer. This may be implemented by expanding the N-way WRR shown in Figure 3a) into an $(N.(2^W-1))$ - way unweighted round robin as shown in Figure 3b), where W equals the number of bits necessary to define the weight, such that if a queue has a weight of w, then it is represented as $(w-1)$ entries in the unweighted round robin list. For example, with 4-bit weights, a 4-way weighted round robin expands to a 60-way unweighted round robin.

[0020] In order to optimize the service intervals to the queues under all weighting conditions, the entries in the unweighted round robin list are distributed such that for each weight

the entries are an equal number of steps apart plus or minus one step. Table 1 below shows an example of such an arrangement of 3-bit weights:

$\underline{w}_n$	$\underline{e}_n$
1	1000000
2	1000100
3	1001010
4	1010101
5	1011011
6	1110111
7	1111111

[0021] In the system described, the arbiter must select one of the nine queues with 4-bit weights, that is 8 virtual output queues as described above and a multicast queue. This expands to a 135-entry unweighted round robin. The implementation of a large unweighted round robin arbiter may be achieved without resorting to a slow iterative shift-and-test method by the technique of "divide and conquer", that is the 135-entry round robin is segmented into 9 sections of 16-entry round robins, each of which may be implemented efficiently with combinational logic (9 x 16 provides up to 144 entries, so that the multicast queue of up to 24 entries may actually be allocated more bandwidth than an individual unicast queue of up to 15 entries).

[0022] Figure 4 illustrates the partitioning of the round robin arbiter. The sorter 41 separates the request vector V (144 bits) into 9 sections of 16-bit vectors , v0 to v8. It also creates nine pointers p0 to p8 for each of the 16-bit round robin blocks 42. The block which corresponds to the existing pointer (which has been saved in register 44) is given a "1" at the corresponding bit location, whilst the other blocks are given dummy pointers initialized to location zero. Each 16-bit round robin block now finds the next "1" in its input vector and outputs its location (g)whether it has to wrap round (w) and whether it has found a "1" in its vector (f). A selector 43 is now able to identify the block which has found the "1" corresponding

to the next "1" in the original 135-bit vector given a signal (s) from the sorter 41. This specifies which round robin block had the original pointer position. The selector 43 is itself a round robin function which may be implemented as a combinational logic function

"find the next block starting at s which has w=false and f=true

(if not found, select s)".

[0023] Figure 5 shows an example of the above process, but for a smaller configuration for clarity. In the example, V = 12 bits, p = 4 bits, v0 - 2 = 2 bits and g0 - 2 = 2 bits. Figure 5 depicts the process performed by Figure 4 and at 51 defines the expanded current pointer (P) and the request vector (V) at 52. The sorter 41 produces segmented vectors (v) and segmented pointers (p) where the blocks marked * are dummies. The segmented results (g) of the round robin are shown at 55 whereas the results of the selector process 43 is shown at 56, defining the expanded next pointer (P).

[0024] The central interconnect 1 provides the cross-connect function in the switch. The bandwidth allocation in the central interconnect is defined by an $(M/N)^2$ -entry central allocation table, which specifies the weights allocated to each possible connection through the central interconnect (the central interconnect has M/N high-bandwidth ports). The central allocation table contains P2 entries, where P=(M/N). Each entry w_{ie} defines the weights allocated to the connection from high-bandwidth port *i* to high-bandwidth port *e*. However, not all combinations of entries constitute a self-consistent set, that is the allocations as seen from the outputs could contradict the allocations as seen from the inputs. A set of allocations is only self-consistent if the sums of weights at each output and input are equal. Figure 6 shows a self-consistent set (a) and a non-self-consistent set (b) of allocation for a 4-port interconnect with 3-bit weights. Inputs are shown at IP and outputs at OP, with the sum designated as Σ. Assuming that the central

allocation table has a self-consistent set of entries, it is possible to define the bandwidth allocation to a link: between input port i and output port a with weight w_{ie} as p_{ie}, where:

$$p_{ie} = \frac{w_{ie}}{\sum_{n=0}^{(p-1)} win}$$

[0025] The egress port table defines how the bandwidth of a high-bandwidth port to the central interconnect 1 is allocated across the virtual output queues. There is no issue with self-consistence as all possible entries are self-consistent. so that the bandwidth allocation for a virtual output queue v with weight w_v is given by:

$$p_f = \frac{w_v}{\sum_{n=0}^{(N-1)} w_n}$$

[0026] Similarly, the ingress port table entries give the bandwidth allocation of a virtual output queue to the ingress ports with weight w_f is given by:

$$p_f = \frac{w_f}{\sum_{n=0}^{(N-1)} w_n}$$

[0027] Therefore the proportion of bandwidth at an egress port v allocated to an ingress port f is given by:

$$p_{fv} = p_f \cdot p_v \cdot p_{ie}$$

[0028] In a switch which is required to maintain strict bandwidth allocation between ports (such as an ATM switch), the tables are set up via a switch management interface from a

connection admission and control processor. When the connection admission and control processor has checked that it has the resources available in the switch to satisfy the connection request, then it can modify the ingress port table, the egress port table and the central allocation table to reflect the new distribution of traffic through the switch.

[0029] In contrast, a switch may be required to provide a "best effort" service. In this case the table entries are derived from a number of local parameters. Two such parameters are the length Iv of the virtual output queue 1~ and the urgency u,. of the virtual output queue. urgency is a parameter which is derived from the headers of the cells entering the queue from the ingress ports.

[0030] A switch may be implemented which can satisfy a range of requirements (including the two above) by defining a weighting function which "mixes" a number of scheduling parameters to generate the table entries in real time according to a set of "sensitivities" to length, urgency and pseudo-static bandwidth allocation. (s_1, sw, ss). The requirement on the function are that it should be fast and efficient, since multiple instances occur in the critical path of a switch. In the system described the weighting function has the form:

$$w_v = \left\{ \frac{1_v^{\,2}}{2^{(1/sl)}} + \frac{p_v}{2^{(1/ss}} + \frac{u_v}{2^{(1/su)}} \right\} . (1 - b_v)$$

where b_v is the backpressure applied from the egress multiplexer,

w_v is the weight of the queue as applied to the scheduler, and

p_v is a pseudo-static bandwidth allocation, such as an egress port table.

[0031] Despite the apparent complexity of this function, it may be implemented exclusively with an adder, multiplexers and small lookup tables , thus meeting the requirement for sped and efficiency. Features of this weighting function are that, for $s_l = 1.0$, $s_s = 0.0$ and $s_u =$

0.0, bandwidth is allocated locally purely on the basis of queue length, with a non-linear function, so that the switch always attempts to avoid queues overflowing. When s1 = 0.0, $s_s = 1.0$ and $s_u = 0.0$, bandwidth is allocated purely on the basis of pseudo-static allocations a described above. Finally, when .$s_l = 0.0$, $s_s = 1.0$ and $s_u = 0.5$, bandwidth is allocated on the basis of pseudo-static allocation but a data source is allowed to "push" some data harder, when the demand arises, by setting the urgency bit in the appropriate cell headers.

[0032] Figure 7 is a block diagram of a small switch based on the above principles, showing the correct number of queues, tables and table entries. In Figure 7 there are two ingress routers 71 and 72, a central cross-bar switch 73 and two egress routers 74 and 75. Each ingress router has two low-bandwidth input ports, A and B for router 71 and ports C and D for router 72. As mentioned previously, each ingress router has an ingress port table such as 77 for router 72 and an egress port table such as 78, whereby the central switch 73 has a central allocation table 79. Assuming that each low-bandwidth port may transport 1Gbps of traffic, each high-bandwidth link may carry 2 Gbps and the switch is required to guarantee the following bandwidth allocations:

Flow bandwidth (Gbps)	Destination Port			
	A	**B**	**C**	**D**
A	0.5	0.1	0.1	0.2
B	0.2	0.2	0.2	0.2
C	-	0.5	-	0.2
D	0.1	0.1	0.6	0.2

then the ingress port table such as 77, egress port table such as 78 and central allocation table 79 would be set up by the connection admission and control processor with the following 4-bit values (note here that there will be rounding errors due to the limited resolution of the 4-bit weights):

Ingress Port Table
(in router 71)

	Source	
	A	**B**
A	15	6
B	3	6
C	3	6
D	6	6

Ingress Port Table
(in router 72)

	Source	
	C	**D**
A	0	3
B	15	3
C	0	13
D	6	5

Ingress Port Table
(in router 71)

	Source
	AB
A	14
B	6
C	6
D	8

Ingress Port Table
(in router 72)

	Source
	CD
A	2
B	12
C	12
D	8

Central Allocation Table

		Destination Router	
		AB	**CB**
Source	**AB**	15	10
	CD	10	15

Claims.

1. A method of scheduling the passage of data cells from M low-bandwidth data sources (4) to M low-bandwidth data destinations (5), said method being performed by a data switching apparatus including:

M/N ingress multiplexers (2; 71, 72), each arranged to receive data cells from a respective set of N said low-bandwidth data sources (4),

MIN egress multiplexers (3; 74, 75), each arranged to transmit data cells to a respective set of N said low-bandwidth data destinations (5),

a master control unit, and

a central switch having M/N high-bandwidth input port,,, arranged to receive data cells from respective said ingress multiplexers (2; 71, 72), and M/N high-bandwidth output ports arranged to transmit data. cells to respective said egress multiplexers (3; 74, 75), the central switch selectively interconnecting the input ports and output ports, under the direction of the master control unit,

the method including:

each said ingress multiplexer (2; 71, 72) maintaining N input queues (22) for queuing data cells received from the N respective said data sources, and maintaining M virtual output queues (24) for queuing data cells directed to respective said data destinations;

and the method being characterized in that:

each ingress multiplexer further maintains a respective ingress port table (27; 77), each ingress port fable having NxM entries, each entry corresponding to a respective combination of a said data source for that ingress port and a said data destination,

each ingress multiplexer transfers data. cells from said input queues to said virtual

output queues with a relative frequency according to value of the corresponding entry of the ingress port table (27; 77);

each ingress multiplexer further maintains a respective egress port table (28; 78), the egress port table, having M entries, each entry corresponding to a respective said data destination,

each ingress multiplier transfers data cells from said virtual output queues to said respective input ports of the central switch with a relative frequency according to the value of the corresponding entry of the egress port table (28; 78);

the master control unit maintains a central allocation table (79) having (M/N)2 entries, each corresponding to a respective combination of an input port and an output port, and

the master control unit controls the central switch to interconnect pairs of said input ports and output ports with a relative frequency according to the value of the corresponding entry of the central allocation table (79);

whereby said ingress port tables (27; 77), egress port tables (28; 78) and central allocation table (79) together determine the bandwidth through. the digital data switching apparatus from each said data source (4) to each said data destination (5).

2. A method according to claim 1 in which each said ingress multiplexer (2; 71, 72), for each virtual output queue (24), transfers data, cells to that virtual output queue from said input queues in accordance with a N-way weighted round robin, using N weights determined respectively by the N entries of the ingress port table for that virtual output queue.

3. A method according to claim 2 in which each weight is defined by a number of bits w, and the N-way weighted round robin for each virtual output queue is implemented by an $N(2^w-1)$-way unweighted round robin using a request vector list constructed by interleaving N

words of (2^W-1) bits each, each word corresponding to a respective input queue and having a number of "1"s determined by the entry of the ingress port table for that input queue and that virtual output queue.

4. A method according to claim 3 in which the request vector list is separated into a plurality of round robin blocks, each corresponding to a respective input queue, a first round robin process being performed independently within each block, and a second round robin process being performed to make a selection among the blocks.

5. A method according to any preceding claim in which the ingress port table (27; 77), the egress port table (28; 78) and the central allocation table (79) are all programmed from an external source.

6. A method according to claim 5 in which the external source uses parameters characterizing the length of each virtual output queue and the urgency of each virtual output queue.

7. A method according to claim 6 in which the eternal source uses a set of sensitivities relating to length, urgency and pseudo-static bandwidth allocation.

8. A digital data switching apparatus for transmitting data from M low-bandwidth data sources (4) to M low-bandwidth data destinations (5), the apparatus including:

M/N ingress multiplexers (2; 71, 72) for receiving data cells from respective sets of N said low-bandwidth data sources (4),

M/N egress multiplexers (3; 74, 75) for transmitting data. cells to respective sets of N said low-bandwidth data destinations (5),

a master control unit, and

a central switch having M/N high-bandwidth input ports arranged to receive data cells from respective said ingress multiplexers (2; 71, 72), and M/N high-bandwidth output ports arranged to transmit data cells to respective said egress multiplexers (3; 74, 75), the central switch being arranged selectively to interconnect the input ports and output ports, under the direction of the master control unit,

each said ingress multiplexer (2; 71, 72) being arranged to maintain N input queues (22) for queuing data cells received from respective said data sources, and to maintain M virtual output queues (24) for queuing data cells directed to respective said data destinations;

characterized in that:

each ingress multiplexer (2; 71, 72) is arranged to maintain a respective ingress port table (27; 77), each ingress port table having NxM entries, each entry corresponding to a respective combination of a said data source and a said data destination, and each ingress multiplexer is arranged to transfer data cells from said input queues to said virtual output queues with a relative frequency according to value of the corresponding entry of the ingress port table;

each ingress multiplexer is arranged to maintain a respective egress port table (28; 78), the egress port table having M entries, each corresponding to a respective said data destination (5), and each ingress multiplier is arranged to transfer data cells from said virtual output queues to said respective input ports of the central switch with a relative frequency according to value of the corresponding entry of the egress port table (28; 78),

and the master control unit is arranged to maintain a central allocation table (79) having $(M/N)^2$ entries, each corresponding to a respective combination of an input port and an output port, and the master control unit controls the central switch to interconnect pairs of said input ports and output ports with a relative frequency according to the value of the corresponding

entry of the central allocation table (79);

whereby said ingress port tables (27; 77), egress port tables (28; 78) and central allocation table (79) together determine the bandwidth through the digital data switching apparatus from each said data source (4) to each said data destination (5).

9. An apparatus according to claim 1 in which each said ingress multiplexer is arranged, for each virtual output queue (24), to transfer data cells to that virtual output queue from said input queues (22) in accordance with a N-way weighted round robin, using N weights determined respectively by the N entries of the ingress port table (27; 77) for that virtual output queue.

10. A apparatus according to claim 11 in which each weight has a number of bits w, and the N-way weighted round robin for each virtual output queue is implemented by an $N(2^w-1)$-way unweighted round robin using a request vector list constructed by interleaving N words of (2^w-1) bits each, each word corresponding to a respective input queue and having a number of "1"s determined by the entry of the ingress port table for that input queue and that virtual output queue.

11. An apparatus according to claim 10 in which the request vector list is separated into a plurality of round robin blocks, each corresponding to a respective input queue, each ingress multiplexer being arranged to preform a first round robin process independently within each block, and a second round robin process to make a selection among the blocks.

12. A apparatus according to any preceding claim further comprising an external. source unit arranged to program the ingress port table (27; 77), the egress port table (28; 78) and the central allocation table (29).

13. An apparatus according to claim 12 in which the external source unit is arranged to operate using parameters characterizing the length of the virtual output queue and the urgency of the virtual output queue.

14. An apparatus according to claim 13 in which the external source unit is arranged to operate using a set of sensitivities relating to the length, urgency and pseudo-static bandwidth allocation.

ABSTRACT

[0033] According to the present invention there is provided a scheduling and arbitration process for use in a digital data switching arrangement of the type in which a central switch under the direction of a master control provides the cross-connections between a number of high-bandwidth ports to which are connected on the ingress side of the central switch a number of ingress multiplexers, one for each high-bandwidth input port and on the egress side a number of egress multiplexers, one for each high-bandwidth output port, each ingress multiplexer including a set of N input queues serving N low-bandwidth data sources and a set of M virtual output queues serving M low-bandwidth output data sources, characterized in that the scheduling and arbitration arrangement includes three bandwidth allocation tables, an ingress port table associated with the input queues and having NxM entries each arranged to define the bandwidth for a particular virtual output queue, an egress port table associated with the virtual output queues and having M entries each arranged to define the bandwidth allocation of a high-bandwidth port of the central switch to a virtual output queue and a central allocation table located in the master control and having $(MxN)^2$ entries each of which specifies the weights allocated to each possible connection through the central switch.

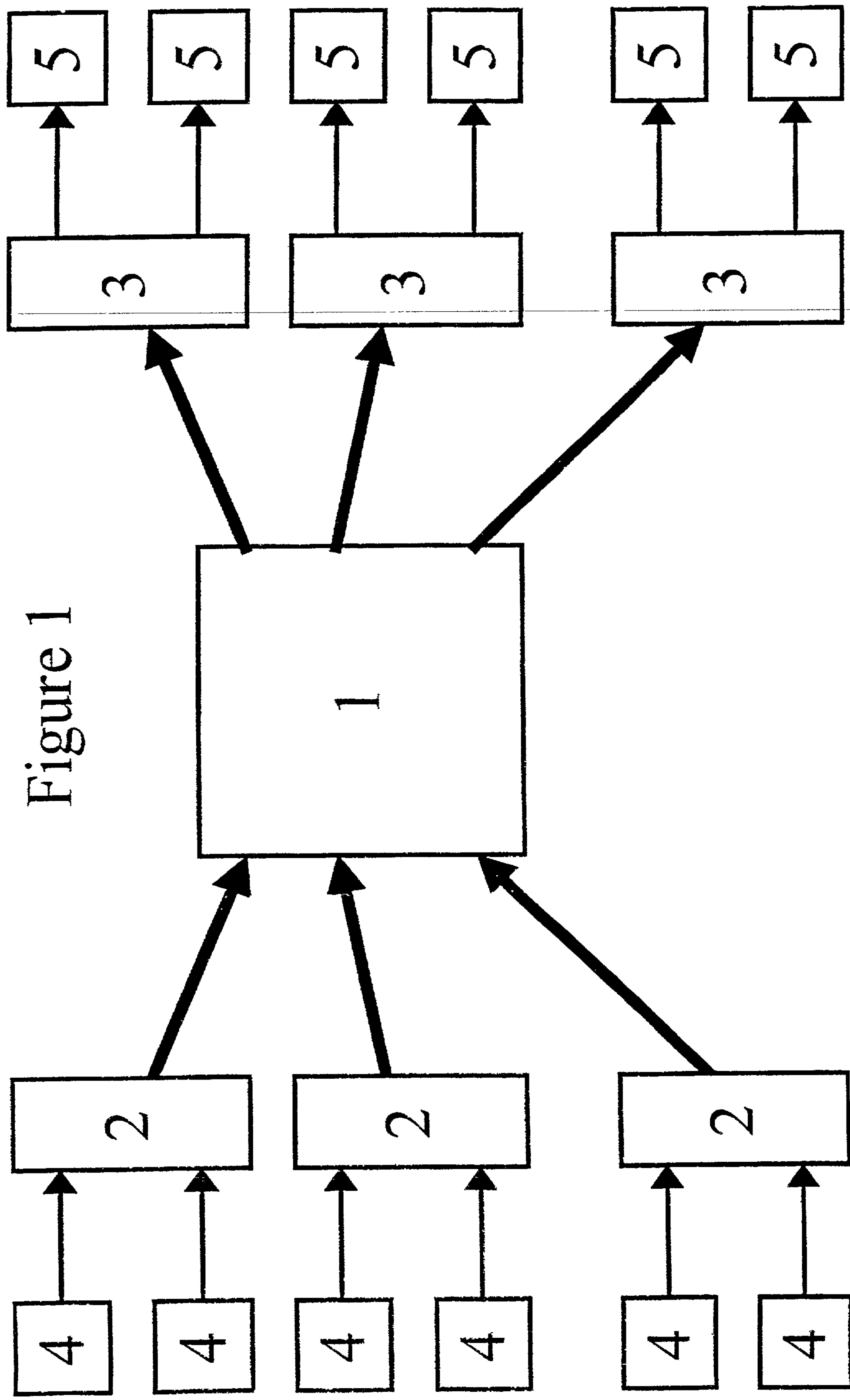


Figure 1

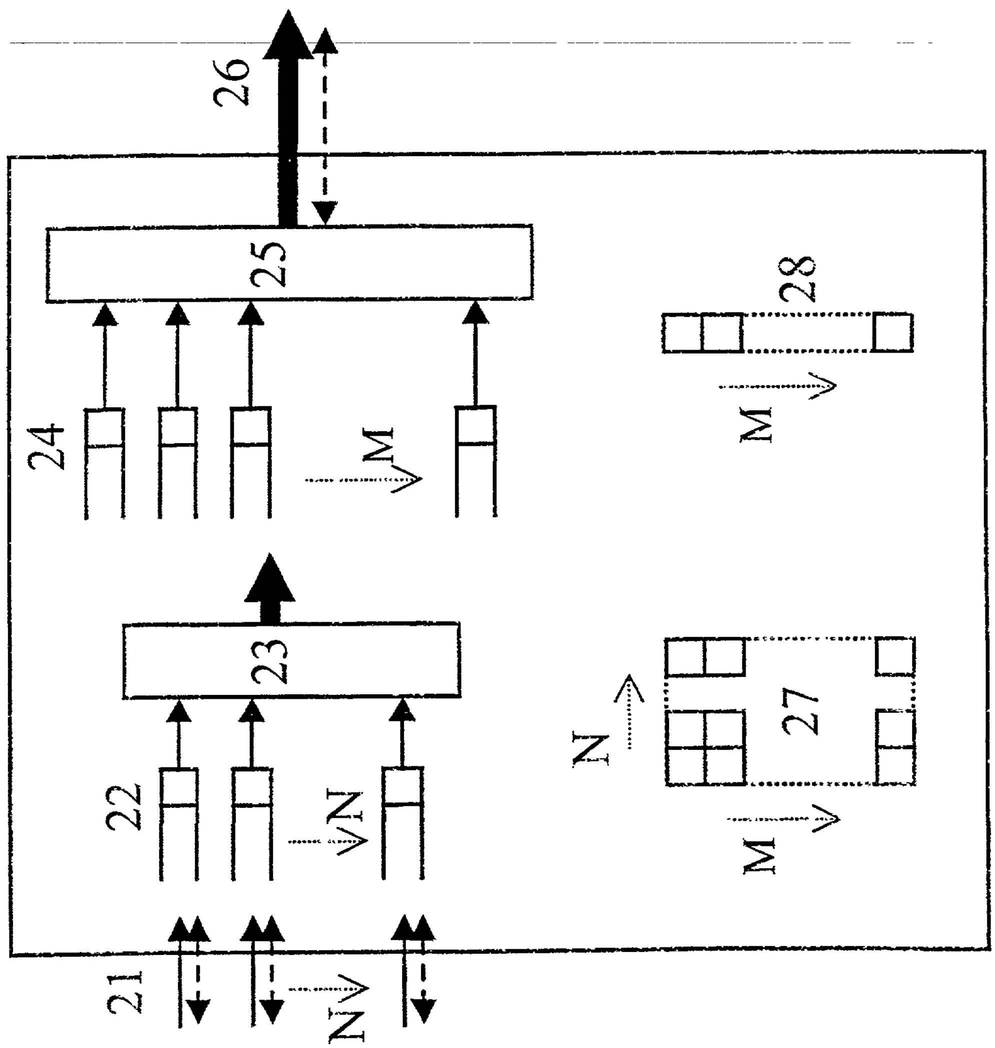


Figure 2

Figure 3

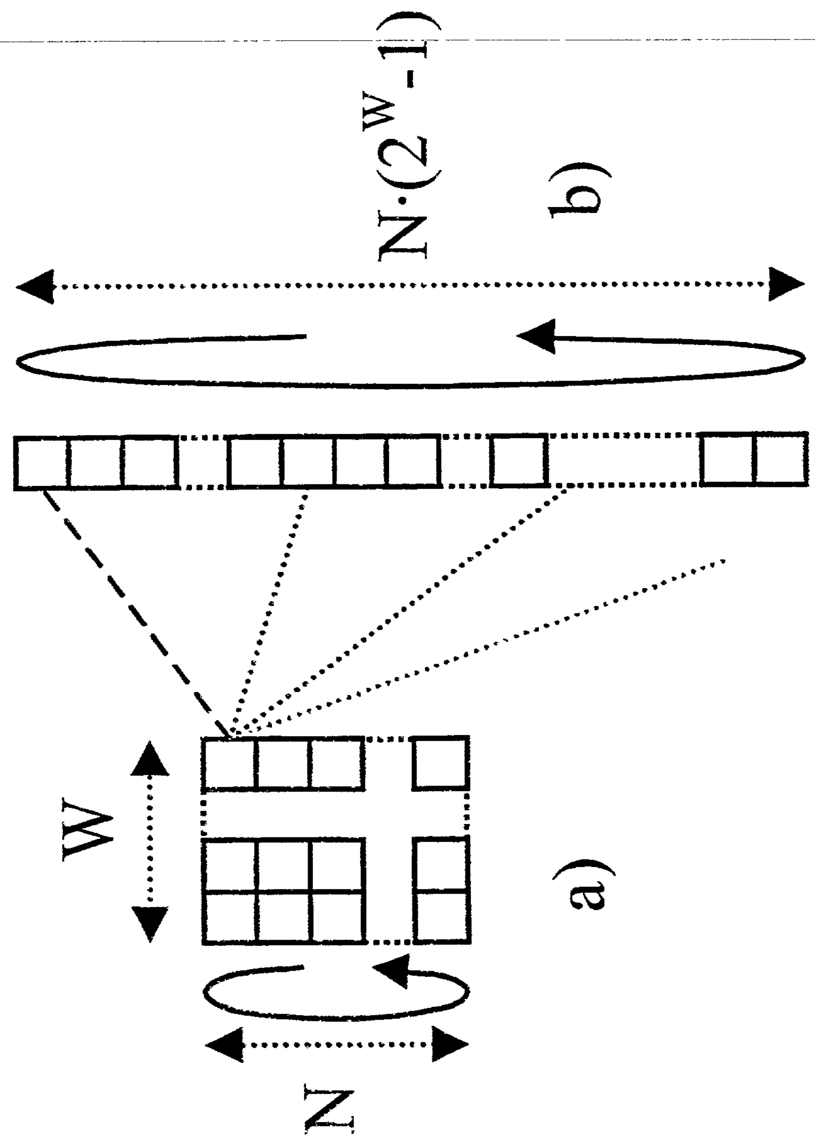


Figure 4

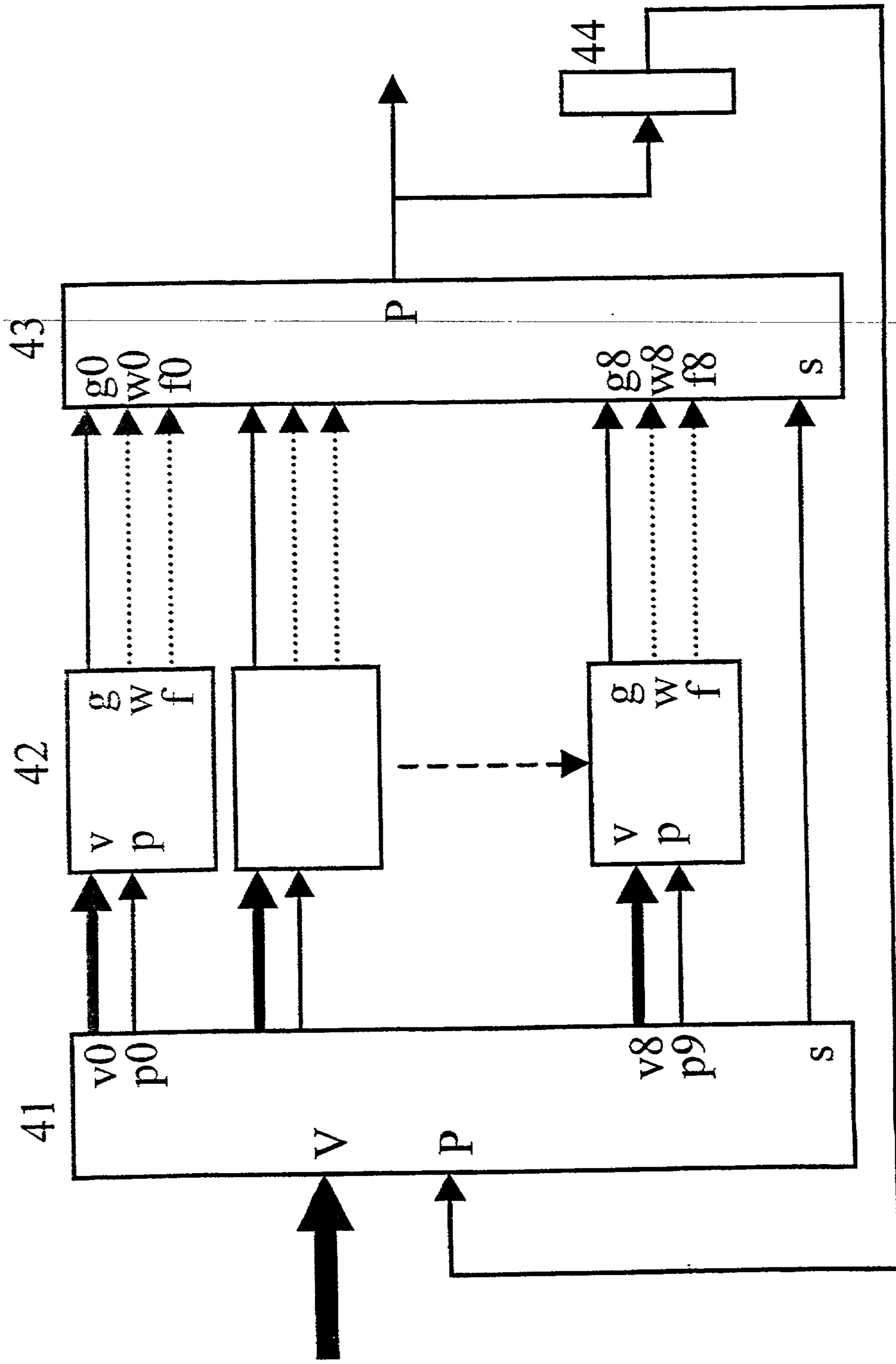
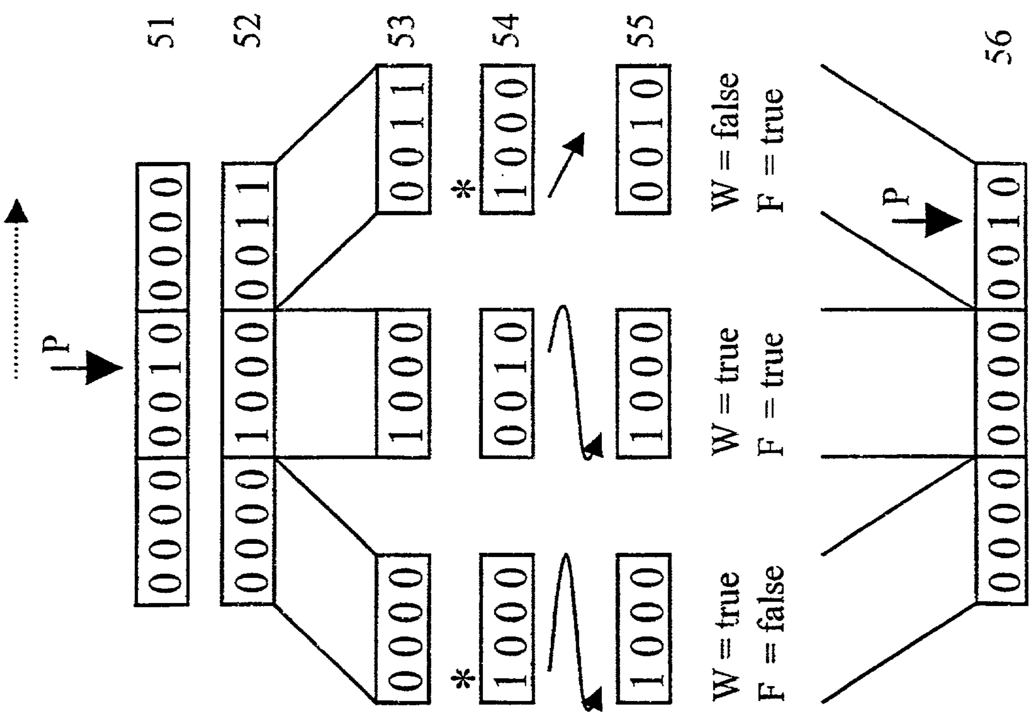


Figure 5



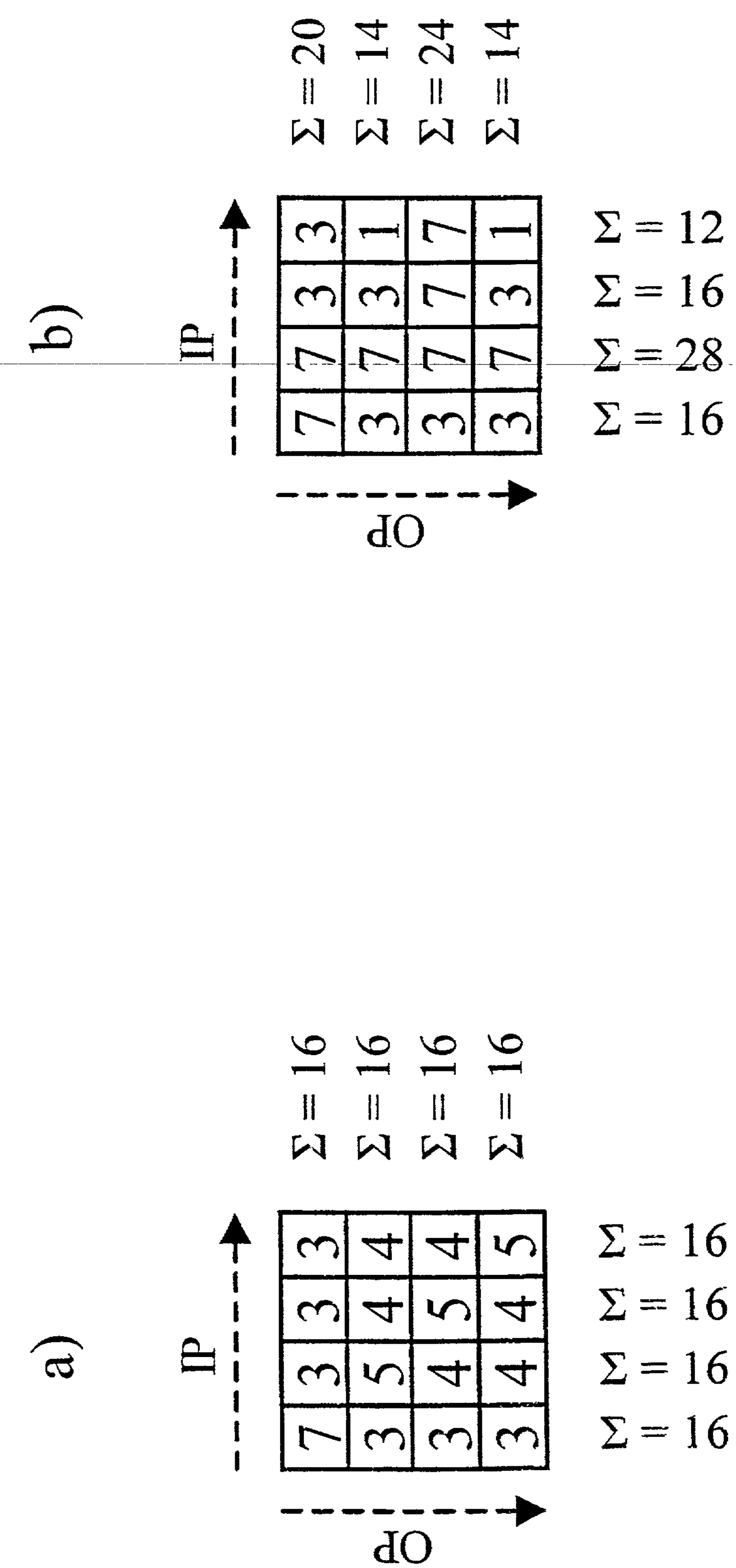

b)
IP
OP
7 7 3 3 Σ = 20
3 7 3 1 Σ = 14
3 7 7 7 Σ = 24
3 7 3 1 Σ = 14
Σ = 16 Σ = 28 Σ = 16 Σ = 12
a)
IP
OP
7 3 3 3 Σ = 16
3 5 4 4 Σ = 16
3 4 5 4 Σ = 16
3 4 4 5 Σ = 16
Σ = 16 Σ = 16 Σ = 16 Σ = 16

Figure 6

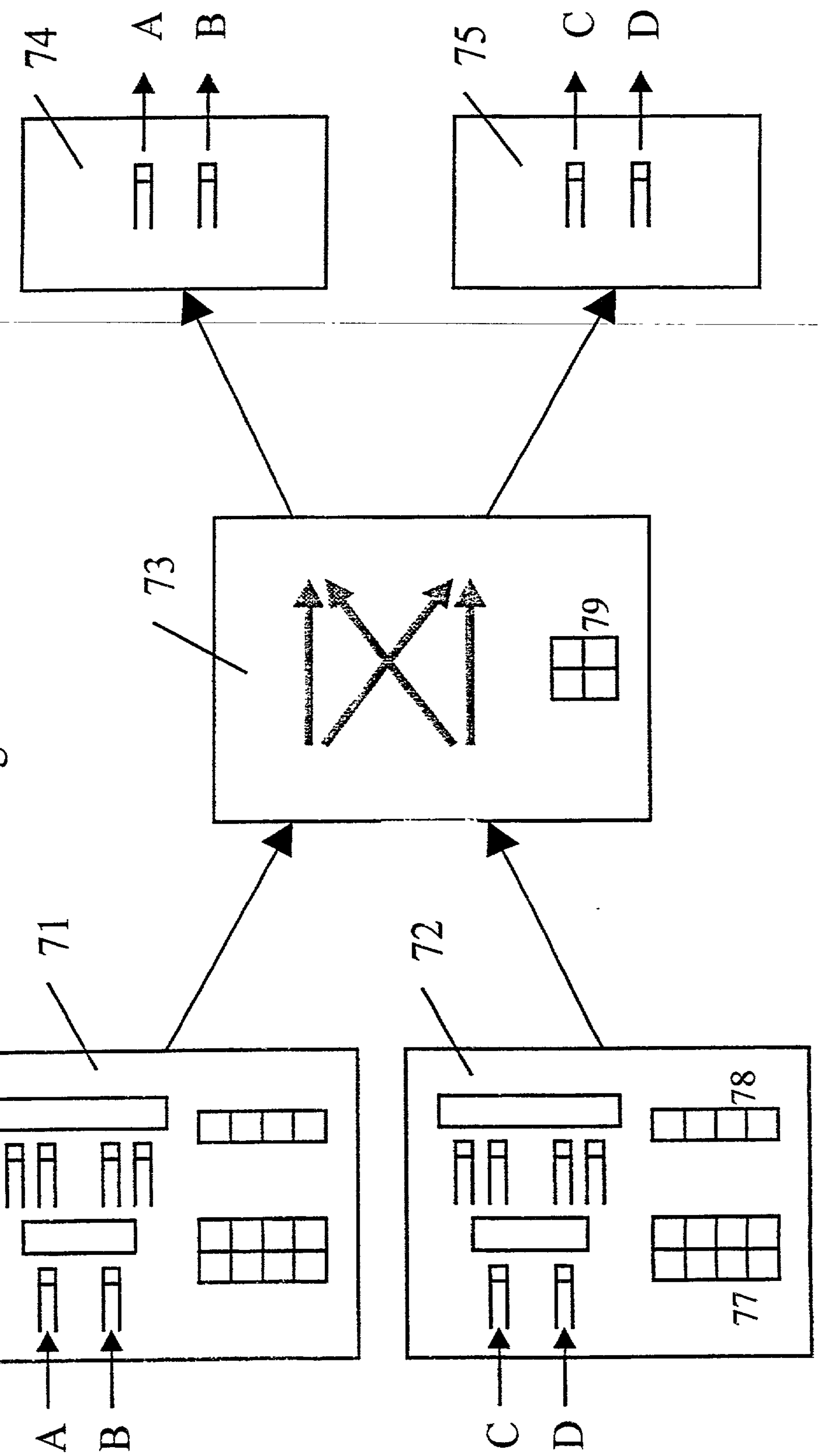


Figure 7

COMBINED DECLARATION AND POWER OF ATTORNEY

(Original, Design, National Stage of PCT, Divisional, Continuation or C-I-P Application)

As a below named inventor, I hereby declare that :
My residence, post office address and citizenship are as stated below next to my name; I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled :

Distributed Hierarchical Scheduling and Arbitration for Bandwidth Allocation

This declaration is of the following type :

[ ]	original	[ ]	divisional
[ ]	design	[ ]	continuation
[x]	national stage of PCT	[ ]	continuation-in-part (C-I-P)

the specification of which : *(complete (a), (b) or (c))*

(a) [ ] is attached hereto.
(b) [ ] was filed on as Application No. and was amended on *(if applicable)*.
(c) [ ] was described and claimed in PCT International Application No. **PCT/GB99/04007** filed on **1 December 1999** and was amended on 27.12.00.

Acknowledgement of Review of Papers and Duty of Candor

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of the subject matter claimed in this application in accordance with Title 37, Code of Federal Regulations § 1.56.

[ ] In compliance with this duty there is attached an information disclosure statement. 37 CFR 1.98.

Priority Claim

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of any PCT International Application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT International Application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application on which priority is claimed

(complete (d) or (e))

(d) [ ] no such applications have been filed
(e) [x] such applications have been filed as follows :

PRIOR FOREIGN/PCT APPLICATION(S) FILED WITHIN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO SAID APPLICATION				
COUNTRY	APPLICATION NO.	DATE OF FILING (day, month, year)	DATE OF ISSUE (day, month, year)	PRIORITY CLAIMED UNDER 35 USC 119
GB	9828143.9	22/12/98		[x] YES NO [ ]
				[ ] YES NO [ ]
				[ ] YES NO [ ]
ALL FOREIGN APPLICATION(S), IF ANY, FILED MORE THAN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO SAID APPLICATION				
				[ ] YES NO [ ]
				[ ] YES NO [ ]
				[ ] YES NO [ ]

Claim for Benefit of Prior U.S.A Provisional Application(s)

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below :

Provisional Application Number	Filing Date

Claim for Benefit of Earlier U.S./PCT Application(s) under 35 U.S.C. 120

(complete this part only if this is a divisional, continuation or C-I-P application)

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) in the manner provided by the first paragraph of Title 35, United States Code § 112, I acknowledge the duty to disclose information as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application :

(Application Serial No.) (Filing Date) (Status) (patented, pending, abandoned)

(Application Serial No.) (Filing Date) (Status) (patented, pending, abandoned)

As a named inventor, I hereby appoint :

Dana M Raymond (18,540) Frederick C Carver (17,021) Francis J Hone (18,662) Joseph D Garon (20,420) Arthur S Tenser (18,839) Ronald B Hildreth (19,498)

Thomas R Nesbitt, Jr (22,075)	Robert Neuner (24,316)	Richard G Berkley
(25,465)		
Richard S Clark (26,154)	Bradley B Geist (27,551)	James J Maune
(26,946)		
John D Murnane (29,836)	Henry Tang (29,705)	Robert C Scheinfeld
(31,300)		
John A Fogarty, Jr (22,348)	Louis S Sorell (32,439)	Rochelle K Seide
(32,300)		

of the firm BAKER & BOTTS, L.L.P., with offices at 30 Rockefeller Plaza, New York, New York 10112, as attorneys to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith

SEND CORRESPONDENCE TO :	DIRECT TELEPHONE CALLS TO :
BAKER & BOTTS, L.L.P. 30 ROCKEFELLER PLAZA, NEW YORK, N.Y. 10112 CUSTOMER NUMBER: 21003	BAKER & BOTTS L.L.P. (212) 705-5000

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR JOINT INVENTOR	LAST NAME JOHNSON	FIRST NAME Ian	MIDDLE NAME David	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS West Kingston Cottage 23 Golden Avenue	CITY East Preston, Littlehampton, Sussex	STATE or COUNTRY Great Britain	ZIP CODE BN16 1QY
DATE MAY 1st 2001	SIGNATURE OF INVENTOR			

FULL NAME OF 2ND JOINT INVENTOR	LAST NAME PIEKARSKI	FIRST NAME Marek	MIDDLE NAME Stephen	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS 20 Gawsworth Road	CITY Macclesfield, Ches	STATE or COUNTRY Great Britain	ZIP CODE SK11 8UE
DATE MAY 2nd 2001	SIGNATURE OF INVENTOR			

FULL NAME OF 3RD JOINT INVENTOR	LAST NAME	FIRST NAME	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE or COUNTRY	ZIP CODE
DATE	SIGNATURE OF INVENTOR			

Applicant or Patentee: Ian David Johnson and Marek Stephen Piekarski
International Appl. No.: PCT/GB99/04007
Filed: 1 December 1999
For: Distributed Hierarchical Scheduling Arbitration for Bandwidth Allocation

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) and 1.27(b)) INDEPENDENT INVENTOR

As a below named inventor, I hereby declare that I qualify as an independent inventor as defined in 37 CFR 1.9 (c) for purposes of paying reduced fees under section 41 (a) and (b) of Title 35, United States Code, to the Patent and Trademark Office with regard to the invention entitled described in:

[ ] the specification filed herewith
[X] International application no.PCT/GB99/04007
[X] filed 1 December 1999

I have not assigned, granted, conveyed or licensed except as shown in the attachment hereto and am under no obligation under contract or law to assign, grant, convey or license, any rights in the invention to any person who could not be classified as an independent inventor under 37 CFR 1.9 (c) if that person had made the invention, or to any concern which would not qualify as a small business concern under 37 CFR 1.9 (d) or a nonprofit organization under 37 CFR 1.9 (e).

Each person, concern or organization to which I have assigned, granted, conveyed, or licensed or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

[ ] no such person, concern, or organization
[X] persons, concerns or organizations listed below*

FULL NAME: Power X Limited
ADDRESS: Stafford Court, 145 Washway Road
Sale,Cheshire M33 7PE, U.K.

[ ] Individual [X] Small Business Concern [ ] Nonprofit Organization

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28 (b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Ian David Johnson

May 1st 2001
Date

Marek Stephen Piekarski

May 2nd 2001
Date

Applicant or Patentee: Ian David Johnson and Marek Stephen Piekarski
Serial or Patent No.: PCT/GB99/04007
Filed or Issued: 1 December 1999
Title: Distributed Hierarchical Scheduling Arbitration for Bandwidth Allocation

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS [37 CFR 1.9 (f) AND 1.27 (c)] - SMALL BUSINESS CONCERN

I hereby declare that I am

[X] the owner of the small business concern identified below :
[ ] an official of the small business concern empowered to act on behalf of the concern identified below :

NAME OF CONCERN Power X Limited

ADDRESS OF CONCERN **145 Washway Road**
Sale
Cheshire M33 7PE, UNITED KINGDOM

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.3-18 and reproduced in 37 CFR 1.9(d), for purpose of paying reduced fees under Section 41(a) and (b) of Title 35, United States Code, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties control or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention, entitled Distributed Hierarchical Scheduling Arbitration for Bandwidth Allocation

by inventor
described in

[ ] the specification filed herewith
[X] the application identified above
[ ] Application Serial No. filed
[ ] Patent No. , issued

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a

small business concern under 37 CFR 1.9(d) or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e). *Note: Separate verified statements are required from each named person, concern, or organization having rights to the invention averring to their status as small entities [37 CFR 1.27].

FULL NAME
ADDRESS
[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

FULL NAME
ADDRESS
[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this Application or Patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate [37 CFR 1.28(b)].

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING ....ANTHONY NEIL PRYCE..............

TITLE OF PERSON OTHER THAN OWNER ....COMPANY SECRETARY.....

ADDRESS OF PERSON SIGNING.22, SILVERDALE, AVE, NEWCASTLE-U-LYME, UK.

SIGNATURE .... DATE...May 1st 2001